Toxicity of Sediments in Northern Puget Sound: A National Perspective

Edward R. Long
National Oceanic and Atmospheric Administration/ORCA

Kenneth A. Dzinbal
Washington Department of Ecology

Introduction

Toxic substances introduced into aquatic ecosystems can bind to particles and collect in deposited sediments, therefore representing a potential toxicological threat the resident biota if concentrations become sufficiently elevated. Toxic chemicals are found in a wide range of concentrations in surficial (recently deposited) sediments of Puget Sound. Although contaminant levels in some areas of Puget Sound have been well characterized with data from many studies, other regions are poorly known. Also, despite source controls initiated in recent decades, some areas remain highly contaminated and thus pose a serious threat to the marine and estuarine ecosystems of Puget Sound.

Despite the availability of data on sediment quality from many studies and regions of Puget Sound, none of these historical data were collected with methods that allowed estimates to be made of the surficial or spatial extent of degradation. Often, studies were performed in the vicinity of specific point sources or other focused areas, thus precluding analyses of the data to determine the actual size and spatial dimensions of the degraded areas.

The National Oceanic and Atmospheric Administration (NOAA) conducts a nationwide program of monitoring and bioeffects assessments via the National Status and Trends (NS&T) Program. NOAA is authorized to conduct this research under Title II of the Marine Protection, Research, and Sanctuaries Act. NOAA has conducted research in numerous bays and estuaries along the Atlantic, Gulf of Mexico, and Pacific coasts in the NS&T Program. In 1997 NOAA elected to include the Puget Sound area in the program. Washington State enacted legislation in 1996 that specifically requires the Puget Sound Water Quality Action Team to ensure continued implementation and coordination of the Puget Sound Ambient Monitoring Program (PSAMP). Through this program, the state is required to monitor and assess the environmental health of Puget Sound. In 1997 NOAA and Washington Department of Ecology entered into the first year of a planned three-year agreement to study adverse biological effects of toxins, such as those found in sediments, in Puget Sound.

The overall goals of this program are to provide information on the presence, severity, and spatial extent of adverse biological effects attributable to toxic chemicals. Data to be generated in this program are intended to be used to estimate the overall environmental health of Puget Sound, to record changes in sediment quality over time as source controls are implemented, to identify areas most in need of source controls and other management actions, and to rank potentially toxic substances of greatest concern. Specific objectives of the program in Puget Sound are:

1. Estimate the spatial extent of chemical contamination, toxicity, and benthic community alterations in surficial sediments;
2. Identify spatial patterns in chemical concentrations, toxicity, and benthic community alterations (possibly leading to the identification of hot spots);
3. Determine the incidence and severity of sediment toxicity;
4. Estimate the apparent relationships between toxicant concentrations and measures of sediment toxicity;
5. Compare and rank the quality of sediment among different regions of Puget Sound; and

6. Determine the temporal trends in contaminant levels and prevalence of liver disease in selected resident demersal fishes.

The purpose of this paper is to outline the methods being used in this program and to document initial results of the toxicity tests performed in northern Puget Sound during 1997. Equivalent research is planned for central Puget Sound in 1998 and for southern Puget Sound in 1999.

Methods

To provide estimates of the spatial extent of sediment degradation, data must be representative of and attributable to the areas in which samples are collected. That is, station location coordinates must be chosen randomly, wherein all candidate longitude/latitude intersections have the same probability of being selected. To provide information on spatial patterns, if any, of sediment degradation, samples must be collected across suspected or known pollution gradients. To determine relationships between measures of contamination and toxicity, chemical analyses and toxicity tests must be performed on portions (aliquots) of the same samples taken synoptically (at the same time). To compare and rank sampling stations based upon a weight of chemical and toxicological evidence, data are needed from a battery of chemical analyses and toxicity tests. The study design selected for this study was intended to satisfy all of these requirements.

The study area chosen extended from the U.S./Canada border south to Port Gardner and Everett Harbor (Figure 1). Emphasis was placed upon four urban areas: Blaine, Bellingham, Anacortes, and Everett where—based upon previous studies—pollution gradients were most likely. Within the selected study area, 33 sampling strata were identified. Strata boundaries were identified as physiographic features expected to have relatively homogenous sedimentological and pollution properties. For example, in bays such as Drayton Harbor and Birch Bay, the strata were identified as a line across the mouth of bay to the 6 ft. isobath. In open-water areas such as the area south of Boundary Bay, the region was arbitrarily sub-divided into roughly equal-sized polygons. Strata were invariably smaller in urban bays where toxicants were expected to occur in heterogeneous or transitional patterns, and, thus, where a denser sampling grid was needed. In areas more distant from urban centers, and, where, therefore, relatively homogeneous, non-polluted conditions were expected, the strata were larger.

Initially, the study plan included Guemes Channel as stratum 20 (Figure 1). However, only rocks and boulders were encountered in this stratum, therefore it was excluded from the study area.

After the dimensions of the strata were determined in discussions between NOAA and Ecology, coordinates of candidate station locations were determined with a computer-aided (NOAA GINPRO) random process. Four alternative locations were determined for each station. The sampling vessel was maneuvered to the first alternative with the aid of differential-corrected GPS. If it was infeasible to sample the first alternative because of obstructions or lack of sediments, the second, third, or fourth alternatives were used.

Sediments were collected in a double van Veen grab sampler following strict criteria for sample acceptance and rejection. Surficial material (upper 2–3 cm) was removed with cleaned high-density plastic scoops and placed in a high-density, polyethylene bucket. The sampler was deployed several times at each station to obtain 7–8 liters of sediment. The composited sample was then homogenized with a paddle. Portions necessary for four toxicity tests and chemical analyses were removed from the composited sample. All sampling equipment was cleaned with ambient seawater, soap, solvent and seawater between stations. Benthic samples were collected with separate deployments of the sampler at each station.

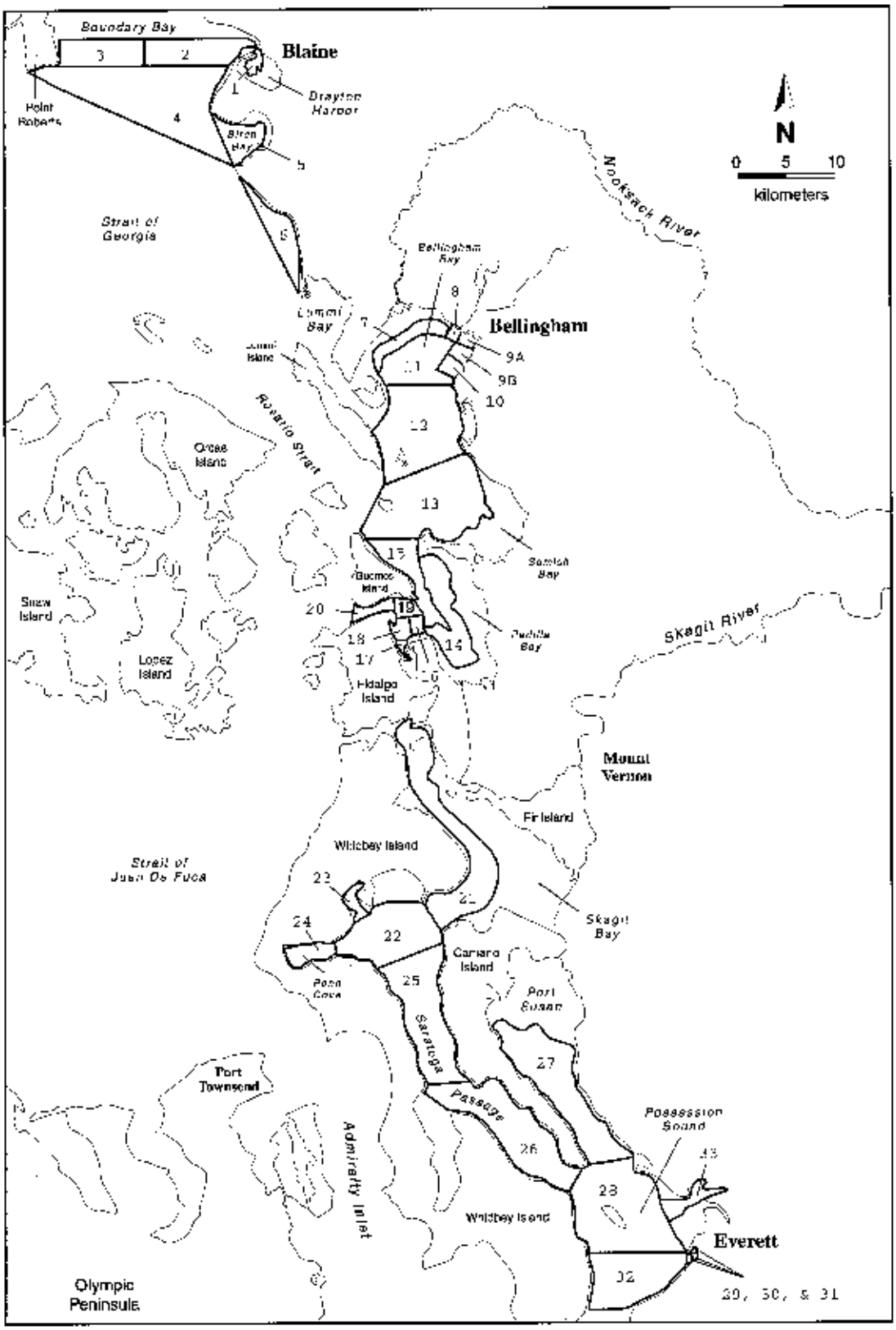

Figure 1. Locations and dimensions of sampling strata within the northern Puget Sound study area.

Toxicity tests were performed with widely accepted protocols. Amphipod survival tests followed protocols of the ASTM (1993), using the species *Ampelisca abdita*. Tests were performed with *A. abdita* in surveys performed in areas along the Atlantic and Gulf of Mexico coasts and as a part of the Environmental Monitoring and Assessment Program (EMAP) estuaries studies (Long et al., 1996). Percent survival in five replicates of 20 animals each was determined after 10-day exposures. Tests of amphipod survival were performed for NOAA by Science Applications International Corporation, as in many previous surveys. This test is regarded as a widely accepted bioassay of relatively unaltered sediments in which acute toxicity is measured with an ecologically-important and relatively sensitive taxonomic group.

Sea urchin fertilization tests followed protocols of the U. S. Geological Survey (USGS) (Carr and Chapman, 1992), using gametes of the purple urchin, *Strongylcentrotus purpuratus*. Pore waters were removed from sediments with gentle pressure, captured in glass, pre-cleaned vials, frozen, thawed and tested in 100%, 50%, and 25% porewater concentrations. Sperm cells were exposed to the porewaters for one hour. Percent fertilization success as determined by the presence of a fertilization membrane around the eggs was determined in five replicates for each sample. These tests were performed by USGS for NOAA, as in numerous other areas. They provide information with a highly sensitive early life stage (sperm cells) exposed to porewaters, the phase in which sediment-associated toxicants are expected to be highly bioavailable.

Cytochrome P-450 assays of the light produced by luciferase in a reporter gene system (RGS) of cultured human liver cells was conducted on all of the samples collected during 1997. In these tests, standard protocols (Anderson et al., 1995, 1996; ASTM, 1997; APHA, 1996) were followed to ensure comparability with data derived from other areas. Approximately 20 g of sediment from each station were extracted with dichloromethane (DCM) to produce 1 mL of DCM containing organic compounds. Small portions of these samples (2–10 µL) were applied to approximately one million human liver cells contained in three replicate wells with 2 mL of culture medium. After 18 hours of incubation, the cells were washed, then lysed, and the solution centrifuged to produce 50 µL extracts. Small portions (10 µL) were used in measures of luminescence. The relative light unit (RLU) from the solvent blank was set equal to unity and all other RLUs were divided by (normalized to) that of the blank. The running average fold induction for 10 nM dioxin is approximately 140 and the induction from 1 µg/mL of benzo(a)pyrene (b[a]p) was 60-fold. Data were converted to mg of b(a)p equivalents per g of sediment. The assay has been performed for NOAA by Columbia Analytical Services, Inc. (CAS) and is responsive to the presence of mixed-function oxidase inducers such as dioxins, furans, high molecular weight PAHs, and coplanar PCBs in tissues and sediments (Anderson et al., 1995).

Microbial bioluminescence (Microtox™) tests were performed with protocols initially developed for Puget Sound (US EPA Region 10, 1990; Schiewe et al., 1985). These tests were run on a portion of the extracts prepared by CAS for the cytochrome P-450 assays. Tests were run in duplicate over a dilution series to determine the EC50 values (the concentrations of sediments that caused a 50% reduction in bioluminescence). USGS in Columbia, MO performed these tests for NOAA, using the same protocols previously used nationwide by USGS and the National Marine Fisheries Service in Charleston, SC. Microtox tests are highly sensitive indicators of the presence of potentially toxic substances in sediments regardless of their bioavailability.

Results of amphipod survival and sea urchin fertilization tests for each sample were compared to those for negative controls to assign statistical significance. In the amphipod tests, sample means were compared to means of tests of a Central Long Island Sound (NY) control previously used in many surveys by SAIC. In the sea urchin tests, sample means were compared to means for controls from Redfish Bay, Texas—an area previously tested by the USGS. Results from the Microtox and cytochrome P-450 tests are currently undergoing review; therefore, calculations of the spatial extent of toxicity are not yet available.

Toxicity data were assigned one of three levels of statistical significance: not significant, significant, or highly significant. When sample means were not significantly different from control means ($p>0.05$), they were classified as not significant (i.e., non-toxic). Samples were classified as significant (i.e., "toxic") when sample means were significantly different ($p<0.05$), but sample means were more than 80% of

control means. Highly toxic samples were those in which sample means were significantly different from controls and less than 80% of controls. The critical value used in the estimates of the spatial extent of toxicity was <80% of controls.

The cumulative frequency distributions of the data were determined by arranging the data in order of descending toxicity. Data from each station were weighted to the sizes of the strata they represented. Spatial extent of toxicity was calculated as the sums of the sample-weighted strata in which the mean toxicity response was less than 80% of controls (Long et al., 1996).

Results

Results of the cytochrome P-450 RGS assay showed relatively non-toxic conditions in most of the 100 samples. The highest responses came from stations clustered within inner Everett Harbor. Figures 2 and 3 show a clear pattern in induction activity among samples collected at stations within and near Everett Harbor. Highest induction activity occurred in the sample from Station 86 (104.6 µg/kg benzo(a)pyrene equivalents) which was collected in Stratum 29 in the inner reaches of the harbor. Toxicity diminished to 33.1 and 52.7 µg/kg B(a)p equivalents in Stations 87 and 88, respectively, within the same stratum, and again to 25–34 µg/kg in samples from Stratum 30, and again to 25–30 µg/kg in samples from Stratum 31 in Port Gardner Bay. Toxicity decreased again among the stations in the outer reaches of Port Gardner Bay (Figure 3).

In the amphipod tests, mean survival ranged from 93% to 99% in the CLIS controls, well within the range of acceptability. Reference toxicant test LC50s ranged between 2.16 to 7.86 mg/L sodium dodecyl sulfate (SDS)—within the acceptable ranged for 10 of 11 test series. Mean survival in test samples, normalized to respective controls, indicated a relatively small range in response—from 82% to >100%.

Mean amphipod survival in sediments from 13 of the 100 stations was significantly different from controls. Thus, the incidence of toxicity was 13%. Three of the stations with significant results were in Strata 2 and 4 south of Boundary Bay, one each was in Strata 10 and 9B in outer Bellingham Bay, one was in Samish Bay, one in outer Padilla Bay, one in Stratum 21 in south Skagit Bay, one in outer Oak Harbor, one in Penn Cove, two in Everett Harbor, and one in the mouth of the Snohomish River.

However, statistical analyses of sample and control means that have very low within-sample variance can show significant differences even when numerical differences between means are very small. The use of the MSD values provides a more rigorous criterion for classifying samples as "toxic," and, therefore, for classifying samples as actually toxic (Thursby et al., 1996). None of the results indicated mean survival was less than 80% of controls. Therefore, the spatial extent of toxicity in the amphipod survival tests was 0% (Table 1).

Table 1. Estimated spatial extent of sediment toxicity in tests of amphipod survival and sea urchin fertilization performed on 100 samples from northern Puget Sound.

Toxicity test	Toxic area (km^2)	Percent of total*
Amphipod survival	0.0	0.0
Urchin fertilization		
• 100% pore water	40.6	5.2
• 50% pore water	11.5	1.5
• 25% pore water	5.9	0.8

* total area: 773.9 km^2

Fertilization success in 100% pore water from the negative controls was 80.6%, 84.6%, and 95.2% in three test batches. Tests of SDS positive controls resulted in EC50s of 2.73–3.11 mg/L SDS. All data were acceptable. In tests of 100% pore water, mean test results normalized to negative controls ranged from 0% fertilization success in several samples to >100%. In tests of 100% pore waters, mean fertilization success was significantly reduced in 15 of the 100 samples; thus, resulting in 15% incidence of toxicity. Based upon the same criteria as in the amphipod tests, the spatial extent of waters (Table 1).toxicity was 5.2% in tests of 100% pore waters, 1.5% in 50% pore waters, and 0.8% in 25% pore waters.

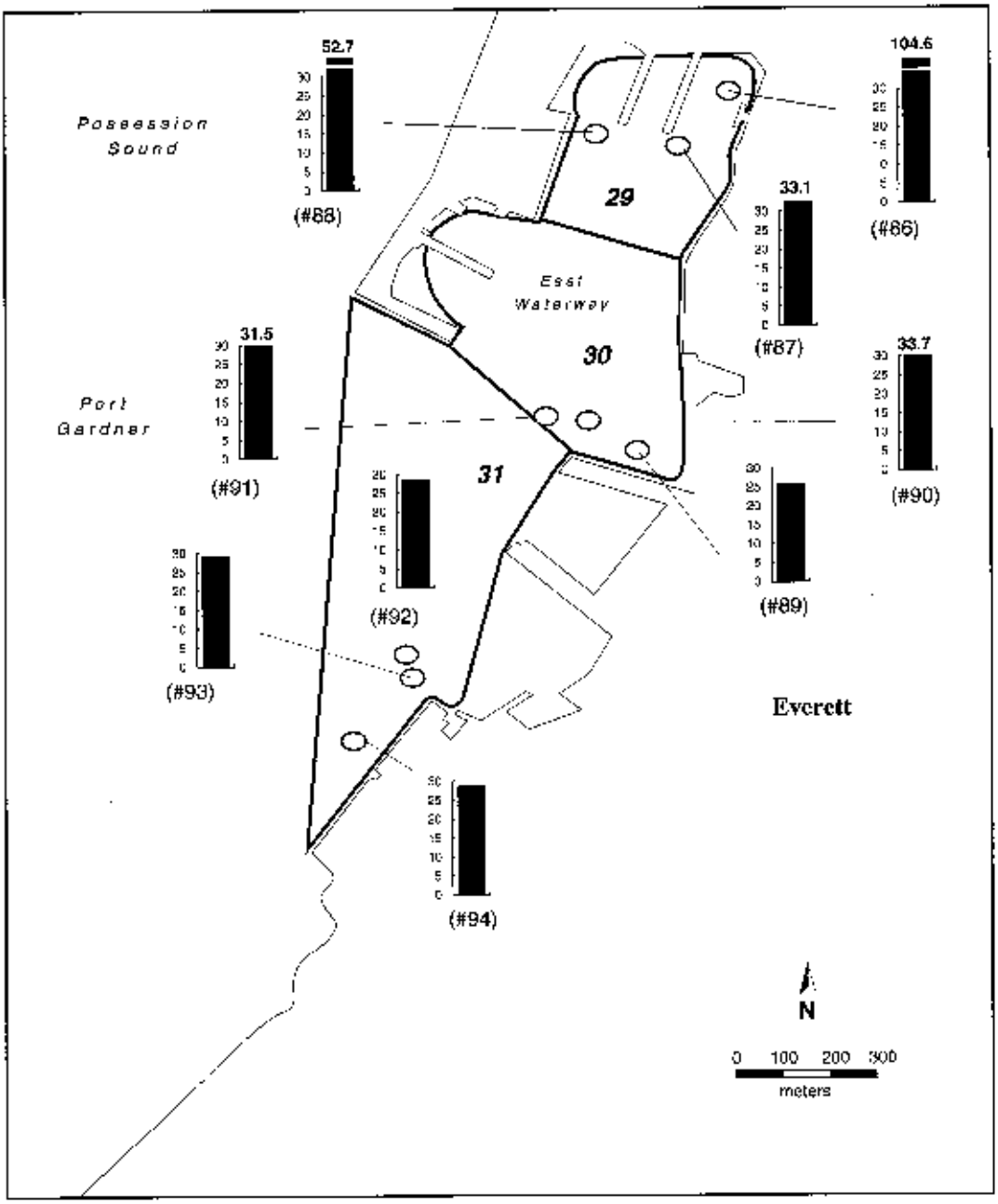

Figure 2. Results of cytochrome P-450 RGS assays of sediments from station in Everett Harbor vicinity (data expressed as µg/kg benzo(a)pyrene equivalents).

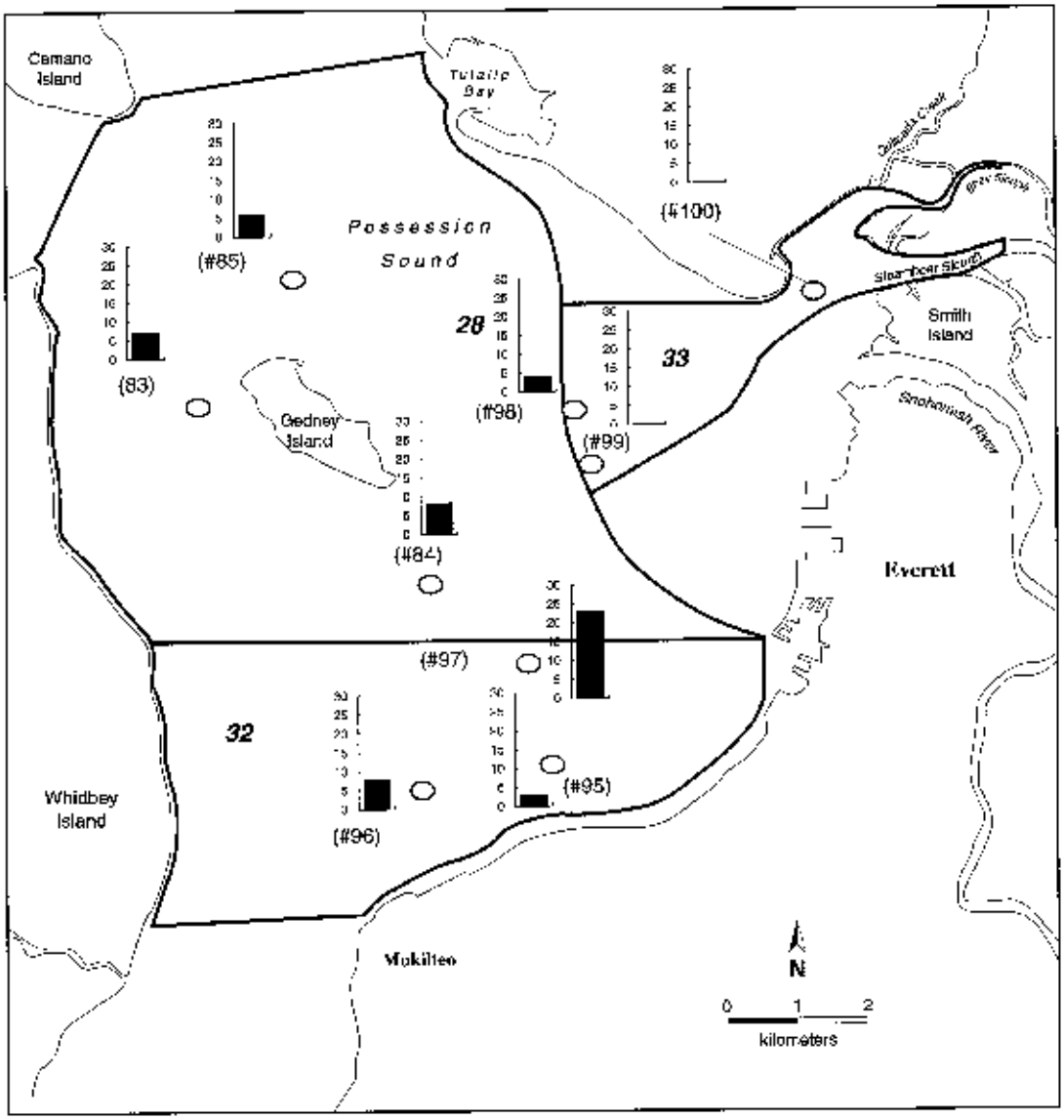

Figure 3. Results of cytochrome P-450 RGS assays of sediments from stations in Port Gardner Bay (data expressed as ¨µg/kg benzo(a)pyrene equivalents).

All nine of the samples from Everett Harbor were highly toxic in tests of 100% pore waters; four of them were highly toxic in tests of both 50% and 25% porewater concentrations. Other samples that were toxic came from Port Susan, Fidalgo Bay, Padilla Bay, Bellingham Bay (one station each), and Drayton Harbor (two stations). Toxicity was most severe in the samples from Drayton Harbor and outer Everett Harbor where fertilization success was 0% to 5% in tests of 100% pore waters.

Discussion and Conclusions

The incidence, severity, and spatial extent of toxicity has been determined in sediments from many different bays and harbors along the Atlantic, Pacific, and Gulf of Mexico coasts (Long et al., 1996). In these surveys, protocols and methods equivalent to those used in Puget Sound were used to ensure comparability of the data. Much of the data was generated by the same laboratories, using the same sources of test animals and controls. To provide perspective to the

results from northern Puget Sound, the data from Puget Sound are compared to those from other areas (Tables 2–3).

Comparable data from amphipod survival tests are available from 24 areas, including northern Puget Sound (Table 2). Except in California where *Rhepoxynius abronius* was used, tests were performed with *Ampelisca abdita*. Surficial extent of toxicity ranged from 0% in many areas to 85% in Newark Bay. The "national averages" calculated with data accumulated after the 1995 surveys (Long et al., 1996) and, again, in 1996 are shown along with the data from each individual area. The samples from northern Puget Sound rank among the least toxic; one of 12 areas with 0% toxicity and well below the national averages of 10.9% in 1995 and 6.9% in 1996. Additional data from US EPA studies performed as a part of the Environmental Monitoring and Assessment Program (EMAP) showed that the surficial areas of toxicity were 0% within the Californian province, 8.4% in the Louisianian province, 2% in the Carolinian province, and 10% in the Virginian province (Long et al., 1996).

Table 2. Spatial extent of toxicity (km^2 and percentages of total area) in amphipod survival tests performed with solid-phase sediments from 24 US bays and estuaries.

				Amphipod survival	
Survey Areas	Year sampled	No. of samples	Total area (km^2)	Toxic area (km^2)	Percent of total area
Newark Bay	93	57	13	10.8	85.0%
San Diego Bay	93	117	40.2	26.3	65.8%
California coastal lagoons	94	30	5	2.9	57.9%
Tijuana River	93	6	0.3	0.18	56.2%
Long Island Sound	91	60	71.86	36.3	50.5%
Hudson-Raritan Estuary	91	117	350	133.3	38.1%
San Pedro Bay	92	105	53.8	7.8	14.5%
Biscayne Bay	95/96	226	484.2	62.3	12.9%
National average: 1995		**1274**	**2532.6**	**277.00**	**10.9%**
Boston Harbor	93	55	56.1	5.7	10.0%
National average: 1996		**1470**	**4158.1**	**286.40**	**6.9%**
Savannah River	94	60	13.12	0.16	1.2%
St. Simons Sound	94	20	24.6	0.10	0.4%
Tampa Bay	92/93	165	550	0.5	0.1%
Galveston Bay	96	75	1351.1	0.0	0.0%
northern Puget Sound	97	100	773.9	0.00	0.0%
Pensacola Bay	93	40	273	0.04	0.0%
Choctawhatchee Bay	94	37	254.47	0	0.0%
Sabine Lake	95	66	245.9	0.00	0.0%
Apalachicola Bay	94	9	187.58	0	0.0%
St. Andrew Bay	93	31	127.2	0	0.0%
Charleston Harbor	93	63	41.1	0	0.0%
Winyah Bay	93	9	7.3	0	0.0%
Mission Bay	93	11	6.1	0.0	0.0%
Leadenwah Creek	93	9	1.69	0	0.0%
San Diego River	93	2	0.5	0.0	0.0%

Sea urchin fertilization tests were performed with *Arbacia punctulata* obtained from the Gulf of Mexico in most areas. In California, gametes of *Strongylocentrotus purpuratus* or embryos of abalone *Haliotis rufescens* were used in the porewater tests. In the sea urchin tests performed with 100% pore waters, northern Puget Sound ranked among the least toxic areas (Table 3). The national averages calculated with data accumulated through 1995 and 1996 were 42.6% and 38.7%, respectively. In

contrast, the results for northern Puget Sound showed the spatial extent of toxicity was about 5.2%. Only three other areas showed less toxicity in this test.

Table 3. Spatial extent of toxicity (km^2 and percentages of total area) in sea urchin fertilization tests performed with 100% sediment pore waters from 23 US bays and estuaries.

				Urchin fertilization @100%	
Survey areas	Year sampled	No. of samples	Total area (km^2)	Toxic area (km^2)	Percent of total area
San Pedro Bay	92	105	53.8	52.6	97.7%
Tampa Bay	92/93	165	550	463.6	84.3%
San Diego Bay	93	117	40.2	25.6	76.0%
Mission Bay	93	11	6.1	4.0	65.9%
Tijuana River	93	6	0.3	0.18	56.2%
San Diego River	93	2	0.5	0.26	52.0%
Biscayne Bay	95/96	226	484.2	229.5	47.4%
Choctawhatchee Bay	94	37	254.47	113.14	44.4%
California coastal lagoons	94	30	5	2.1	42.7%
National average: 1995		**940**	**2082.6**	**886.3**	**42.6%**
Winyah Bay	93	9	7.3	3.1	42.2%
National average: 1996		**1136**	**3723.26**	**1439.73**	**38.7%**
Apalachicola Bay	94	9	187.58	63.6	33.9%
Galveston Bay	96	75	1351.1	432.0	32.0%
Charleston Harbor	93	63	41.1	12.5	30.4%
Savannah River	94	60	13.12	2.42	18.4%
Boston Harbor	93	55	56.1	3.8	6.6%
Sabine Lake	95	66	245.9	14.0	5.7%
Pensacola Bay	93	40	273	14.4	5.3%
northern Puget Sound	97	100	773.9	40.6	5.2%
St. Simons Sound	94	20	24.6	0.65	2.6%
St. Andrew Bay	93	31	127.2	2.28	1.8%
Leadenwah Creek	93	9	1.69	0	0.0%

In conclusion, data available thus far from the tests suggest that toxicity is not very severe or widespread in northern Puget Sound sediments. The surficial area in which acute toxicity to adult crustaceans was observed represented 0% of the total area, and the extent of toxicity in the sublethal fertilization test was very small—5% of the total. Cytochrome P-450 RGS assays showed highest induction mainly in samples from inner Everett Harbor. Microbial bioluminescence appeared to be most affected also in samples from Everett Harbor. Results from all four tests indicate that toxicity was most severe in inner Everett Harbor, and to a lesser extent in a few other urban bays such as Drayton Harbor and Bellingham Bay. Data from chemical analyses will be used to identify substances that may have contributed to toxicity. Data from benthic community analyses will be examined to determine if toxicity observed in laboratory tests is also expressed *in situ* among resident biota. Collectively, data from the chemical, toxicological, and benthic analyses will be compiled to provide an overall synopsis of sediment quality in northern Puget Sound.

References

Anderson, J. W., S. S. Rossi, R. H. Tukey, T. Vu, and L. C. Quattrochi. 1995. A biomarker, 450 RGS, for assessing the potential toxicity of organic compounds in environmental samples. Environmental Toxicology and Chemistry 7 (140): 1159–1169.

APHA. 1996. P450 reporter gene response to dioxin-like organics. pp. 24–25. Method 8070. In: Standard Methods for the Examination of Water and Wastewater, 19th Edition Supplement. American Public Health Association. Washington, D.C.

ASTM. 1993. Standard guide for conducting solidphase, 10-day, static sediment toxicity tests with marine and estuarine infaunal amphipods. ASTM E 1367-92. American Society for Testing Materials. Philadelphia, PA. 24 pp.

ASTM. 1997. E 1853-96. Standard guide for measuring the presence of planar organic compounds which induce CYP1A, reporter gene test system. pp. 1392–1397. American Society for Testing and Materials. Philadelphia, PA.

Carr, R.S. and D. C. Chapman. 1992. Comparison of solid-phase and pore water approaches for assessing the quality of marine and estuarine sediments. Chemistry and Ecology 7: 19–30.

Long, E. R., A. Robertson, D. A. Wolfe, J. Hameedi, and G. M. Sloane. 1996. Estimates of the spatial extent of sediment toxicity in major U. S. estuaries. Environmental Science & Technology 30 (12): 3585–3592.

Schiewe, M. H., E. G. Hawk, D. I. Actor, and M. M. Krahn. 1985. Use of a bacterial bioluminescence assay to assess toxicity of contaminated marine sediments. Canadian Journal of Fisheries and Aquatic Science 42: 1244–1248.

Thursby, G. B., J. Heltshe, and K. J. Scott. 1997. Revised approach to toxicity test acceptability criteria using a statistical performance assessment. Environmental Toxicology and Chemistry 16 (6): 1322–1329.

U.S. EPA Region 10. 1990. Recommended protocols for conducting laboratory bioassays on Puget Sound sediments. Revised, April, 1990. Puget Sound Estuary Program, US Environmental Protection Agency. Seattle, WA. 79 pp.